

09/487,401
10/7/8

DIALOG(R)File 347:JAPIO
(c) 1999 JPO & JAPIO. All rts. reserv.

BEST AVAILABLE COPY

03340146 **Image available**
LOG RECORD CONTROL SYSTEM

PUB. NO.: 03-003046 [JP 3003046 A]
PUBLISHED: January 09, 1991 (19910109)
INVENTOR(s): AOKI HISANOBU
KATO MASAMICHI

APPLICANT(s): HITACHI LTD [000510] (A Japanese Company or Corporation), JP
(Japan)
HITACHI MICRO COMPUT ENG LTD [470864] (A Japanese Company or
Corporation), JP (Japan)

APPL. NO.: 01-136039 [JP 89136039]

FILED: May 31, 1989 (19890531)

ABSTRACT

PURPOSE: To speed up restoring operation from a transaction fault by preparing an area to be used between processes in common, and when the *reading*/writing of *log* information from/in a common memory is controlled or the common memory is filled with data, setting up a *log* control manager to be written in a secondary storage.

CONSTITUTION: The *log* record control system consists of a *data* *base* control system 12, a *log* control manager 13, a *data* *base* 14, a *log* file 15, user processes 16, 17, and common memories 18, 19 corresponding to user tasks. The *log* control manager 13 writes *requested* *log* information in the common memory, and when the common memory is filled with data, writes out the information to the secondary storage medium. When a transaction fault occurs, a data recovery *request* is outputted to the

manager 13, and when the *log* information concerned exists, restoring processing is immediately executed. When the information does not exist, the *log* information group stored in the secondary storage is retrieved and the *log* information group concerned is written in a buffer to execute recovery processing. Consequently, data recovery processing at the time of the occurrence of the transaction fault can be rapidly executed.

⑬ 日本国特許庁(JP)

⑭ 特許出願公開

⑫ 公開特許公報(A) 平3-3046

⑮ Int.Cl.⁹

G 06 F 12/00
11/34
12/00

識別記号

3 0 2 N
A
3 0 2 A

庁内整理番号

8944-5B
7343-5B
8944-5B

⑯ 公開 平成3年(1991)1月9日

審査請求 未請求 請求項の数 2 (全5頁)

⑰ 発明の名称 ログ記録管理方式

⑱ 特 願 平1-136039

⑲ 出 願 平1(1989)5月31日

⑳ 発 明 者 青 木 久 延 東京都小平市上水本町5丁目22番1号 日立マイクロコンピュータエンジニアリング株式会社内

㉑ 発 明 者 加 藤 正 道 神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

㉒ 出 願 人 株式会社日立製作所 東京都千代田区神田駿河台4丁目6番地

㉓ 出 願 人 日立マイクロコンピュータエンジニアリング株式会社 東京都小平市上水本町5丁目22番1号

㉔ 代 理 人 弁理士 小川 勝男 外1名

明 細 書

1. 発明の名称

ログ記録管理方式

2. 特許請求の範囲

1. トランザクション障害時にトランザクション

によつて更新されたデータを障害前の状態に復元させるトランザクション処理システム及び各プロセスが自分のアドレス空間の一部をプロセス間で共有できる領域(共有メモリと称す)を持つことができるシステムにおいて、共有メモリに一時的にトランザクションによつて更新されたデータ情報(ログ情報と称す)を積み書きできる管理マネージャを設けることによりトランザクション障害時のデータ回復を円滑に行なわせるログ記録管理方式。

2. 請求項1記載の方式において、共有メモリが満杯になつた場合には二次記憶装置に格納を行い、その際に格納したログ情報を一括して取りだせるようにログ情報の前後にデータサイズを格納するエリアと二次記憶に書き出した順番を

格納する順番エリアを設けたことを特徴とするログ記録管理方式。

3. 発明の詳細な説明

(産業上の利用分野)

本発明はトランザクション障害が発生した際にトランザクションによつて変更されたデータを実行前の状態に円滑に復元させるためのログの記録管理方式に関する。

(従来の技術)

従来より、更新したファイルを元に戻すために更新前の状態をログ情報としてファイルとは別の媒体に記録することにより、異常時にファイルの回復を行う方法が知られている「上田恭雄著、「オンラインシステム」、明発堂、pp.93~98」。しかし、上記回復に際して該当ログ情報を記録媒体から効率よくサーチしてデータの復元を行う方法については述べられていなかった。

(発明が解決しようとする課題)

上記従来技術は、障害時のデータ回復処理のために、障害を大きく3つに分けてそれぞれ対処し

ている。

(1) トランザクション障害

ユーザの誤操作、アプリケーションプログラムのエラーやデッドロックのために発生するひとつのトランザクション（データ操作に関するひとまとまりの処理）に閉じた障害である。この場合トランザクション開始時点に更新前ログを使ってデータベース（データ）の状態を復元する。

(2) システム障害

システムプログラム（オペレーティングシステム、データベース管理システム等）のエラーや電圧異常により発生する、データベースシステム全体にわたる障害である。この場合の回復処理は、システムの再立ちあげ後、上記（1）の処理と同様な処理を行う。

(3) メディア障害

データベースを格納している二次記憶媒体（ハードディスク、磁気テープ等）の障害である。この場合の回復処理は、データベースのバックアップダンプ及びログにより行う。

上記目的を達成するために本発明においては、トランザクションを発行する各プロセスごとに自分のアドレス空間の一部にプロセス間で共有できる領域（共有メモリと称す）を持たせ、共有メモリへのログ情報の読み書きの管理、及び共有メモリが満杯時に二次記憶に共有メモリの内容を書き出すログ管理マネージャを設定する。これにより、障害発生時に該当するプロセスの共有メモリに残されたログ情報をもとに障害回復をおこなう。また、共有メモリにログ情報が存在しない場合には、二次記憶に格納されたログファイルを検索する。ファイルに格納されたログ情報は、各プロセス対応のかたまりとして格納される（共有メモリの内容を二次記憶に書き出すための）データ回復のための一連のログ情報を高速に取り出すことができる。

〔作用〕

トランザクションを要求する各プロセスに共有メモリを持たせる。各プロセスより実行されたトランザクションのログ情報が、ログ管理マネージャ

大型計算機（ホスト）では、多数ユーザがデータベースを使用することからデータの復元を煩雑しているため、前記（1）の様な、もつとも頻りに発生する障害からの復元処理の高速性に欠けている。すなわち、従来技術はログ情報を直接二次記憶媒体に格納しており、ここに格納されるログ情報は複数のプロセス（ユーザ）から要求されるため、時間的順序に従って二次記憶に書き込まれる。このため、同一トランザクションのログ情報の記録が連続していない。従ってトランザクション障害時にその回復を行う場合に、トランザクションが行なった一連の操作を記録したログ情報を検索する場合に1個ずつランダムに二次記憶媒体をアクセスしながら復元処理を行うため、高速な復元処理の点について配慮されておらず、復元処理のための操作時間に問題があった。

本発明は、トランザクション障害からの復元操作を高速にすることを目的としてなされたものである。

〔問題を解決するための手段〕

ログ管理マネージャは要求されたログ情報を要求をおこなったプロセスの共有メモリに書き込む。共有メモリが満杯の場合には、その情報（ログ情報群と称す）を二次記憶媒体に書き出す。

これにより二次記憶にはログ情報が順々に分類されずに同一のトランザクションによるログ情報が固まって格納される。

トランザクション障害が発生するとデータの回復作業がログ管理マネージャに要求され、ログ管理マネージャは該当するトランザクションから属するプロセスの共有メモリを参照し、該当ログ情報が存在しているかをチェックする。存在していれば、その情報をもとにただちに復元処理を行う。存在していなければ、二次記憶に格納されたログ情報群を検索し、該当するログ情報群をログ管理マネージャが管理しているバッファに読みこんで、回復処理を行う。

以上によりトランザクション障害時のデータの回復処理を高速に行うことができる。

〔実施例〕

以下、本発明の一実施例を第1図～第5図により説明する。

第1図はデータベースシステム11の全体構成図であり、データベース管理システム（以降DBMSと称す）12、ログ管理マネージャ13、データベース14、ログファイル15、ユーザプロセス16、17とユーザタスク対応の共有メモリ18、19から成る。

第2図は共有メモリ18、19及びログファイル15に記録されるログ情報の形式20、27である。トランザクション識別子21はデータベース14の操作を行ったトランザクションを識別するものである。動作識別子23はトランザクションの動作を識別するもので、開始、終了、データの書き込み、削除から成る24、25はデータを改めた場合にその前後の値が格納される。26には、トランザクションの開始から終了までのシーケンシャル番号が格納される。22（22A、22B）は1図に格納されるログ情報長である。

する。この情報は、ログ情報を該当するユーザの共有メモリから抽出する際に使用される。

- (2) ユーザプロセス16がデータベースの更新をDBMS12に要求するとデータの更新情報がログ管理マネージャ13に送信（34）される。
- (3) ログ管理マネージャ13はDBMS12より送られた情報をもとにログ情報20を作成する。次に共有メモリのアドレス管理テーブル32を参照してログ情報を記録する共有メモリ18のアドレスを求めて（アドレスのサーチはユーザプロセスの識別子をキーにして行う）該当共有メモリ18にログ情報を書き込む35、共有メモリに情報を記録するための空き領域がない場合には、すでに書き込まれている共有メモリ18の内容をログファイル15に書き出した後（36）、書き込み要求のあったログ情報を共有メモリに書き込む（35）。

次にトランザクション終了時の回復操作を第4図を用いて説明する。

- (1) ユーザプロセス16がトランザクション終了

ここで、データ長22-Aと22-Bは、ログファイル15にログ情報を書き出した際に、ログ情報をファイルの先頭からも最後からも取り出すことを可能とするために設定したものである。

27はログファイル15に格納されるログ情報の形式である。28は共有メモリに1図に格納されるログ情報20のN個分（ログ情報群と称す）含むログ情報の長さである。29には同一の共有メモリからログファイルに格納された順順番号が格納される。なおログ情報長28が前後に付いているのはログファイルを先頭及び最後からアクセスを可能にするためである。

以下ログ記録の処理手順を第3図を用いて説明する。

- (1) ユーザプロセス16が起動されるとユーザ用の共有メモリ18が確保され、共有メモリの先頭アドレスとユーザの識別子がログ管理マネージャ13に送信される31。ログ管理マネージャはこの情報を自分の持つ管理テーブル32（共有メモリのアドレス管理テーブル）に格納

を引き続き、ユーザプロセス16は、ログ管理マネージャ13にトランザクションによつて変更されたデータの回復を要求（41）する。

- (2) ログ管理マネージャ13は、アドレス管理テーブル32より、要求があつたユーザプロセス16の共有メモリ18のアドレスを求めることにより共有メモリ18の参照（42）を行う。

以下回復処理を第5図のフローチャートを用いて説明する。

トランザクションの終了時を基準にそれに一番近いトランザクションのログ記録が共有メモリ18中に存在するかをチェック（51）する。存在すればその情報をもとにデータ回復を開始（52）、次に近いトランザクションのログ記録を共有メモリ18からサーチして、存在すれば同様にデータの回復を行う。この一連の操作をトランザクションの開始まで続ける。開始の判断はログ記録エントリ、動作識別子23で行う。共有メモリ18中に該当する情報が無ければログファイル15より該当するログ情報群27をログ管理マネージャの

バッファに読み込み(43、56、57)バッファの内容をもとにサーチ(51)して回復処理(52、53)を行う。

以上によりトランザクション障害時の障害回復を高速に行うことができる。

〔発明の効果〕

本発明の如く、ユーザのデータ操作ミス、アプリケーションエラー、デッドロック等データベースの障害中で最も頻繁に発生すると考えられるトランザクション障害に対して、ユーザプロセスが保持する共有メモリログ管理マネージャによつてログ情報の読み書きを行わせることにより、障害時にデータ回復を高速に行うことができる。また共有メモリが満杯の場合には、共有メモリ単位に同一トランザクションのログ情報のかたまりを二次記憶媒体に格納できるので、トランザクション障害時に該当するログ情報が共有メモリ中に存在しない場合でも、二次記憶から該当するログ情報のかたまりとして一度にログ管理マネージャの管理するバッファに読み込んで回復処理が行えるの

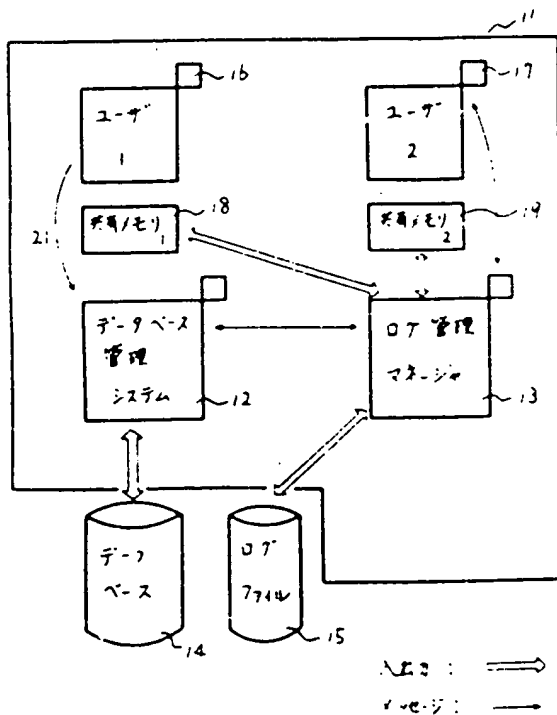
で、二次記憶からログ情報をサーチするための入出力時間を短縮でき、高速なデータ回復処理を行うことができる。

4. 図面の簡単な説明

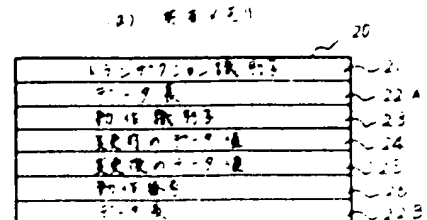
第1図は本発明の一実施例のシステムの全体構成図、第2図はログ情報を共有メモリ及びバッファに格納するためのデータ形式図、第3図は、ログ情報の格納のための手順説明図、第4図はトランザクション障害時のデータ回復の手順説明図、第5図はトランザクション障害時のデータ回復のための詳細フローチャート図である。

代理人 弁護士 小川 隆

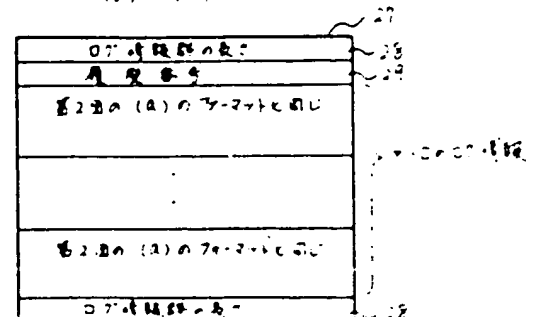
第 1 図
データベースシステムの全体構成



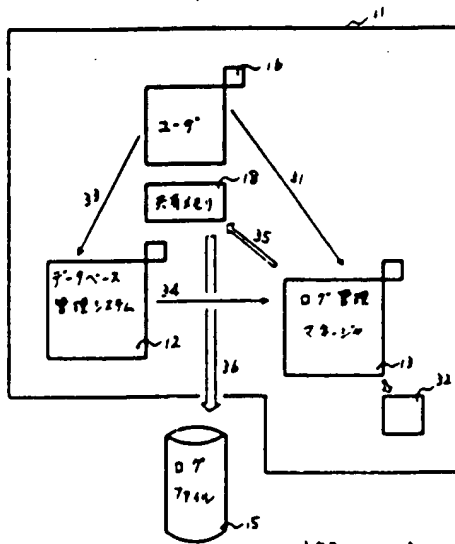
第 2 図
ログ情報の格納形式



(ロ) ログファイル

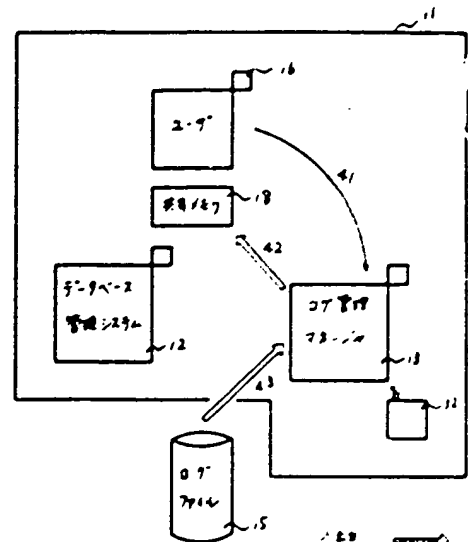


第 3 図
ログ情報の書き込み処理



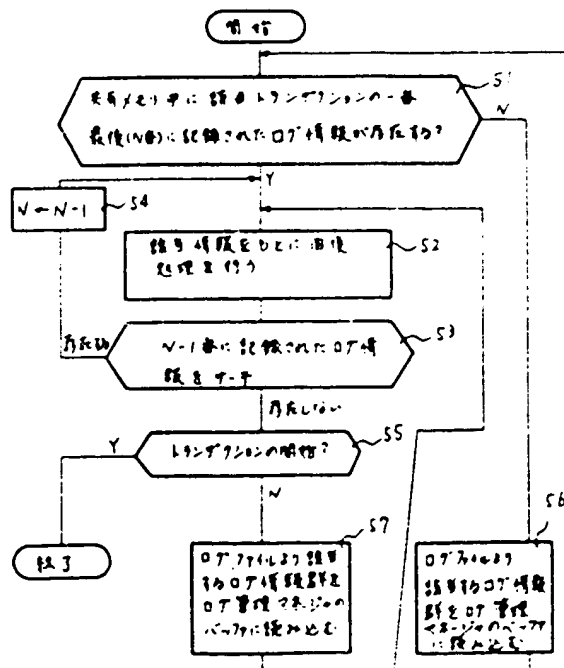
11 ホストコンピュータ
31 共有メモリ上のログ管理モジュール

第 4 図
トランザクション発生時の回復処理



11 ホストコンピュータ
32 共有メモリ上のログ管理モジュール

第 5 図
トランザクション発生時の回復処理



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)